Properties of a Five-Band Matrix and Its Application to Boundary Value Problems

RIAZ A. USMANI
University of Manitoba
Winnipeg, Manitoba, Canada

Communicated by Richard Varga

1. INTRODUCTION

Consider the two-point boundary value problem

$$y''(x) = f(x)y(x) + g(x), \qquad y(a) = y_a, \qquad y(b) = y_b, \tag{1.1}$$

where $f(x)$ and $g(x) \in C$ over the interval $[a, b]$, and a, b, y_a, and y_b are real finite constants with $a < b$. The numerical solution of the two-point boundary value problem (1.1) is most commonly obtained by finite difference methods. We introduce a finite set of grid points

$$x_i = a + ih \qquad (i = 0, 1, \ldots, N + 1),$$

where $x_0 = a$, $x_{N+1} = b$, and $h = (b - a)/(N + 1)$. In defining any finite difference analog for the numerical solution of (1.1) we require that the discretization error $e_i = y(x_i) - y_i$, i.e., the difference between the exact solution $y(x_i)$ of problem (1.1) at the grid point x_i and its approximate value y_i obtained by solving the finite difference equations can be made arbitrarily small as the mesh size h tends to zero. Also we need a bound on the discretization error e_i. Such an error bound is quite useful in the selection of h. Both Henrici [4] and Varga [6] have considered problem (1.1) with $f(x) \geqslant 0$ on $[a, b]$. Varga proved (1962) that the resulting error is $0(h^2)$ based on finite difference equations (Δ.E.'s).

$$y_n - 2y_{n+1} + y_{n+2} = h^2 y''_{n+1} \qquad (n = 0, 1, \ldots, N - 1). \tag{1.2}$$

Linear Algebra and Its Applications 4(1971), 101–115

The inequality derived by Varga is

$$\|e\|^* = \max_i |e_i| \leqslant \frac{h^2(b-a)^2 M_4}{96}, \tag{1.3}$$

where $M_i = \max_{a \leqslant x \leqslant b} |d^i y/dx^i|$, and $e = (e_i)$, the error vector; see [6, Th. 6.2, p. 165].

A five-point difference analog was considered (1964) by Bramble and Hubbard [2]. They approximated problem (1.1) by the following Δ.E.'s.

(i) $y_{n-1} - 2y_n + y_{n+1} = h^2 y_n''$,

truncation error (T.E.) $t_n = \dfrac{h^4}{12} y^{(4)}(\xi_n) \qquad (n = 1, N)$,

(ii) $\frac{1}{12}[-(y_{n-2} + y_{n+2}) + 16(y_{n-1} + y_{n+1}) - 30y_n] = h^2 y_n''$,

$$\text{T.E.} \quad t_n = -\frac{h^6}{90} y^{(6)}(\xi_n) \qquad (n = 2, 3, \ldots, N-1), \tag{1.4}$$

or, in matrix form, $Ay = d$, where A is a five-band matrix of order N. Both $y = (y_i)$ and $d = (d_i)$ are column vectors. The error bound derived by Bramble and Hubbard was

$$\|e\| \leqslant h^4 \left[\tfrac{1}{2} M_4 + \frac{(b-a)^2}{720} M_6 \right]. \tag{1.5}$$

Recently the error bounds obtained by Varga [6], by Henrici [4], and by many others, e.g., [1, 5], with $f(x) \geqslant 0$ on $[a, b]$ have been greatly improved by Fischer and Usmani [3].

The above is a simple example of the way certain band matrices arise in the numerical solution of boundary value problems. The purpose of this paper is to derive some properties of a five-band matrix and show how the results may be applied in the error analysis of the two-point boundary value problem (1.1). The author relies entirely on the theory of Δ.E.'s for generating the elements of the inverse of the five-band matrix considered.

* For a vector $v = (v_i)$, $|v| = (|v_i|)$, $\|v\| = \max_i |v_i|$, and for a matrix $A = (a_{ij})$, $\|A\| = \max_i \sum_j |a_{ij}|$.

2. FIVE-BAND MATRIX AND ITS PROPERTIES

Let $J = (j_{mn})$ be a five-band matrix of order N given by

$$J = \begin{bmatrix} 2c-1 & 2-c & -1 & & & \\ 2-c & 2c-2 & 2-c & -1 & & \\ -1 & 2-c & 2c-2 & 2-c & -1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & & -1 & 2-c & 2c-2 & 2-c \\ & & & -1 & 2-c & 2c-1 \end{bmatrix}, \tag{2.1}$$

where c is a real parameter.

Also let $P = (p_{mn})$ be a tridiagonal matrix of order n, with elements

$$p_{mn} = \begin{cases} c > 2, & m = n, \\ -1, & |m-n| = 1; \qquad m, n = 1, 2, \ldots, n, \\ 0, & \text{otherwise}. \end{cases} \tag{2.2}$$

Define D_n = determinant of the matrix P. Then, on expanding the determinant D_n by Laplace's method across the first row for $n \geq 3$, we obtain the Δ.E.

$$D_n = cD_{n-1} - D_{n-2}. \tag{2.3}$$

If we define $D_{-1} = 0$ and $D_0 = 1$, then the Δ.E. (2.3) is also valid for $n = 1$ and 2. The solution of the Δ.E. (2.3) along with initial conditions

$$D_{-1} = 0, \qquad D_0 = 1$$

is given by

$$D_n = \frac{\sinh(n+1)\theta}{\sinh\theta}, \qquad 2\cosh\theta = c; \tag{2.4}$$

see [3, p. 129]. The elements of $J^{-1} = (a_{mn})$ will be expressed in terms of the polynomials D_n. We will now derive several results concerning the matrix J.

Reducible and Monotone Matrices. Let W be the set of first n integers, $W = \{1, 2, \ldots, n\}$. A matrix $M = (m_{ij})$ of order n is said to be reducible if it is possible to decompose W into two nonempty, disjoint subsets S

and T, such that $m_{ij} = 0$ for $i \in S$ and $j \in T$. A matrix which is not reducible is called irreducible.

A square matrix $M = (m_{ij})$ of order n with real elements is called monotone if $Az \geqslant 0$ implies $z \geqslant 0$, where $z = (z_i)$ is a column vector with n components. We shall use the following results concerning monotone matrices in establishing the theorems that will follow subsequently.

(i) A monotone matrix is nonsingular;

(ii) a matrix M is monotone if the elements of the inverse matrix M^{-1} are nonnegative;

(iii) let the matrix $M = (m_{ij})$ be irreducible and satisfy the conditions

$$m_{ij} \leqslant 0, \qquad i \neq j \qquad (i, j = 1, 2, \ldots, n),$$

and

$$\sum_{j=1}^{n} m_{ij} \quad \begin{cases} \geqslant 0 & (i = 1, 2, \ldots, n), \\ > 0, & \text{for at least one } i; \end{cases}$$

then M is a monotone matrix;

(iv) let the two matrices M_1 and M_2 be monotone such that $M_1 \geqslant M_2$; then $M_2^{-1} \geqslant M_1^{-1}$. (2.5)

For proof of these results, see [4, p. 360].

THEOREM 2.1. *The matrix J is irreducible if $c \neq 2$.*

Proof. The proof follows immediately from a well-known theorem of graph theory [6, Th. 1.6], namely, a complex matrix $A = (a_{ij})$ of order n is irreducible if its directed graph $G(A)$ is strongly connected. Since $c \neq 2$, $j_{mn} \neq 0$, $|m - n| = 1$; $m, n = 1, 2, \ldots, N$. Now it is an easy matter to verify that the directed graph $G(J)$ is strongly connected, hence J is irreducible. To complete the proof of the theorem we must prove that, for $c = 2$, J is reducible. To prove this we rely on the definition of a reducible matrix given above. Take $W = \{1, 2, \ldots, N\}$, $S = \{1, 3, 5, \ldots\}$, the subset of W consisting of odd integers and $T = \{2, 4, 6, \ldots\}$, the subset of W consisting of even integers. Clearly, $j_{mn} = 0$ for $m \in S$ and $n \in T$, since the elements of the matrix J for $c = 2$ are given as

$$j_{mn} = \begin{cases} 5, & m = n = 1 \quad \text{or} \quad N, \\ 4, & m = n \qquad (m, n = 2, 3, \ldots, N-1), \\ -1, & |m - n| = 2, \\ 0, & \text{otherwise}; \end{cases}$$

hence J is reducible.

THEOREM 2.2. *The matrix J is a monotone matrix if $c > 2$.*

Proof. The proof follows immediately from (2.5(iii)) since, for $c > 2$, the off-diagonal elements of J are negative and the sums of the elements in any row of J are nonnegative and positive for the 1st, 2nd, $(N-1)$st, and Nth rows.

We now determine the elements of $J^{-1} = (a_{mn})$ explicitly. On multiplying the rows of J by the nth column of J^{-1}, we obtain the following Δ.E.'s $(n = 2, 3, \ldots, N-1)$.

$$\begin{aligned}
\text{(i)} \quad & -a_{m-2,n} + (2-c)a_{m-1,n} + (2c-2)a_{m,n} + (2-c)a_{m+1,n} \\
& -a_{m+2,n} = 0 \qquad (m = 2, \ldots, n-1), \qquad \text{with} \quad a_{0,n} = 0, \\
& (2c-1)a_{1,n} + (2-c)a_{2,n} - a_{3,n} = 0; \\
\text{(ii)} \quad & -a_{n-2,n} + (2-c)a_{n-1,n} + (2c-2)a_{n,n} + (2-c)a_{n+1,n} \\
& -a_{n+2,n} = 1 \qquad (n \neq 1, N); \\
\text{(iii)} \quad & -a_{m-2,n} + (2-c)a_{m-1,n} + (2c-2)a_{m,n} + (2-c)a_{m+1,n} \\
& -a_{m+2,n} = 0 \qquad (m = n+1, \ldots, N-1), \qquad \text{with} \quad a_{N+1,n} = 0, \\
& -a_{N-2,n} + (2-c)a_{N-1,n} + (2c-1)a_{N,n} = 0.
\end{aligned} \tag{2.6}$$

The solution of the Δ.E. (2.6(i)) of order 4 with the associated two initial conditions can easily be obtained in the form

$$a_{mn} = mA + B(\alpha^m - \beta^m), \qquad m = 1, 2, \ldots, n-1, n, n+1, \tag{2.7}$$

where A and B are arbitrary constants to be determined later and α, β are the roots of the quadratic equation

$$x^2 + cx + 1 = 0, \qquad c > 2;$$

that is, $\alpha, \beta = (-c \pm \sqrt{c^2 - 4})/2 = -e^{\theta}, -e^{-\theta}, 2\cosh\theta = c.$

Similarly the solution of Δ.E. (2.6(iii)) with the associated conditions is

$$a_{mn} = -(N - m + 1)C + \frac{D}{\alpha^{N+1}}(\alpha^{N-m+1} - \beta^{N-m+1}),$$

$$m = n - 1, n, n + 1, \ldots, N, \tag{2.8}$$

where C and D are arbitrary constants.

The elements $a_{n-1,n}, a_{n,n}, a_{n+1,n}$ can be obtained either from (2.7) or (2.8). On equating the expressions for a_{mn}, $m = n - 1, n, n + 1$, obtained from (2.7) and (2.8), respectively, we obtain three equations in unknowns A, B, C, and D. The fourth equation is obtained on substituting in (2.6(ii)) the values of a_{mn}, $m = n - 2, n - 1, n, n + 1, n + 2$, derived from (2.7) and (2.8). Thus the unknowns A, B, C, D are determined from the equations

$$(n - 1)A + B(\alpha^{n-1} - \beta^{n-1}) + (N - n + 2)C$$

$$- \frac{D}{\alpha^{N+1}}(\alpha^{N-n+2} - \beta^{N-n+2}) = 0,$$

$$nA + B(\alpha^n - \beta^n) + (N - n + 1)C - \frac{D}{\alpha^{N+1}}(\alpha^{N-n+1} - \beta^{N-n+1}) = 0,$$

$$(n + 1)A + B(\alpha^{n+1} - \beta^{n-1}) + (N - n)C - \frac{D}{\alpha^{N+1}}(\alpha^{N-n} - \beta^{N-n}) = 0,$$

$$(n + 2)A + B[(c^2 - 2)(\alpha^n - \beta^n) - (\alpha^{n-2} - \beta^{n-2})] + (N - n - 1)C$$

$$- \frac{D}{\alpha^{N+1}}(\alpha^{N-n-1} - \beta^{N-n-1}) = 1. \tag{2.9}$$

On solving the system (2.9), we obtain

$$A = \frac{N - n + 1}{(c + 2)(N + 1)},$$

$$B = -\frac{\alpha^{N-n+1} - \beta^{N-n+1}}{(c + 2)(\alpha^{N+2} + \beta^{N+2} - \alpha^N - \beta^N)},$$

$$C = \frac{-n}{(c + 2)(N + 1)},$$

$$D = \frac{(\alpha^n - \beta^n)\alpha^{N+1}}{(c+2)(\alpha^{N+2} + \beta^{N+2} - \alpha^N - \beta^N)}. \tag{2.10}$$

On substituting the values of A, B in (2.7) and of C, D in (2.8), respectively, we obtain

$$a_{mn} = \begin{cases} \dfrac{m(N-n+1)}{(c+2)(N+1)} - \dfrac{(\alpha^m - \beta^m)(\alpha^{N-n+1} - \beta^{N-n+1})}{(c+2)(\alpha^{N+2} + \beta^{N+2} - \alpha^N - \beta^N)}, & m \leqslant n, \\ \dfrac{n(N-m+1)}{(c+2)(N+1)} - \dfrac{(\alpha^n - \beta^n)(\alpha^{N-m+1} - \beta^{N-m+1})}{(c+2)(\alpha^{N+2} + \beta^{N+2} - \alpha^N - \beta^N)}, & m \geqslant n. \end{cases} \tag{2.11}$$

Noting that

$$\alpha = -e^\theta, \qquad \beta = -e^{-\theta},$$

$$\alpha^j + \beta^j = 2(-1)^j \cosh j\theta,$$

$$\alpha^j - \beta^j = 2(-1)^j \sinh j\theta,$$

and making use of the identities

$$\sinh C + \sinh D = 2 \sinh \frac{C+D}{2} \cosh \frac{C-D}{2},$$

$$\cosh C - \cosh D = 2 \sinh \frac{C+D}{2} \sinh \frac{C-D}{2},$$

and (2.4), we further obtain

$$a_{mn} = \begin{cases} \dfrac{m(N-n+1)}{(c+2)(N+1)} + (-1)^{n-m} \dfrac{D_{m-1}D_{N-n}}{(c+2)D_N}, & m \leqslant n \\ \dfrac{n(N-m+1)}{(c+2)(N+1)} + (-1)^{m-n} \dfrac{D_{n-1}D_{N-m}}{(c+2)D_N}, & m \geqslant n \end{cases} \quad (n \neq 1, N). \tag{2.12}$$

For the determination of the first column of J^{-1} we proceed as follows. The elements of the first column of J^{-1}, namely, a_{m1}, $1 \leqslant m \leqslant N$, satisfy the following Δ.E.

$$-a_{m-2,1} + (2-c)a_{m-1,1} + (2c-2)a_{m1} + (2-c)a_{m+1,1} - a_{m+2,1} = 0$$

$$(m = 2, 3, \ldots, N-1), \qquad \text{with conditions}$$

$$a_{0,1} = 0,$$

$$(2c-1)a_{11} + (2-c)a_{21} - a_{31} = 1,$$

$$-a_{N-2,1} + (2-c)a_{N-1,1} + (2c-1)a_{N,1} = 0,$$

$$a_{N+1,1} = 0. \tag{2.13}$$

The solution of the Δ.E. in (2.13) with the associated conditions is

$$a_{m1} = [(N-m+1)/(N+1) + (-1)^{m-1}D_{N-m}/D_N]/(c+2),$$

$$1 \leqslant m \leqslant N. \tag{2.14}$$

The Nth column of J^{-1} can be determined by symmetry; in fact,

$$a_{mN} = a_{N-m+1,1}, \qquad 1 \leqslant m \leqslant N.$$

On combining (2.12) and (2.14) and noting that J is symmetric, monotone matrix, i.e., $a_{mn} \geqslant 0$ for all m, n, we summarize the preceding results in Theorem 2.3.

THEOREM 2.3. *Let the matrix $J = (j_{mn})$, $c > 2$; then $J^{-1} = (a_{mn})$ is symmetric and*

$$a_{mn} = (n(N-m+1)/(N+1) + (-1)^{m-n}D_{n-1}D_{N-m}/D_N)/(c+2) \geqslant 0,$$

$$m \geqslant n.$$

3. APPLICATIONS

The properties of the matrix J derived in Section 2 will now be applied to obtain numerical solution and perform error analysis of the two-point boundary value problem (1.1).

We will study the following Δ.E.'s, given by (3.1), for approximating the boundary value problem (1.1) with $f(x) \geqslant 0$ on $[a, b]$.

$$\text{(i)} \quad -cy_0 + (2c-1)y_1 + (2-c)y_2 - y_3 + h^2 \sum_{i=0}^{3} \bar{\beta}_i y_i'' = 0,$$

$$\text{(ii)} \quad -(y_n + y_{n+4}) + (2-c)(y_{n+1} + y_{n+3}) + (2c-2)y_{n+2}$$

$$+ h^2 \sum_{i=0}^{4} \beta_i y_{n+i}'' = 0 \qquad (n = 0, 1, \ldots, N-3),$$

$$\text{(iii)}\quad -y_{N-2}+(2-c)y_{N-1}+(2c-1)y_N-cy_{N+1}$$
$$+h^2\sum_{i=0}^{3}\bar{\beta}_{3-i}y''_{N-i+2}=0, \tag{3.1}$$

where $y_i''=f_iy_i+g_i$, $y_i\simeq y(x_i)$, the exact solution of problem (1.1) at the point $x=x_i$, $f_i\equiv f(x_i)$, $g_i\equiv g(x_i)$, $i=0,1,2,\ldots,(N+1)$, and

$$\beta_0=\beta_4=(-c+18)/240,$$
$$\beta_1=\beta_3=(12c+104)/120,$$
$$\beta_2=(97c+14)/120,$$
$$\bar{\beta}_0=c/12,\qquad \bar{\beta}_1=(10c+1)/12,\qquad \bar{\beta}_2=(c+10)/12,\quad\text{and}$$
$$\bar{\beta}_3=1/12. \tag{3.2}$$

The local T.E.'s associated with (3.1(i)) and (3.1(iii)) are given by

$$t_n=-\frac{c+1}{240}h^6y^{(6)}(\xi_n),\qquad c\neq -1\qquad (n=1,N),$$

where $x_0<\xi_1<x_3$ and $x_{N-2}<\xi_N<x_{N+1}$, and the local T.E. associated with (3.1(ii)) is

$$t_n=\frac{31c-190}{60{,}480}h^8y^{(8)}(\xi_n),\qquad c\neq\frac{190}{31}\qquad (n=2,3,\ldots,N-1),$$

where $x_{n-2}<\xi_n<x_{n+2}$, provided

$$2<c\leqslant 14/3. \tag{3.3}$$

The details of the derivation of the Δ.E.'s (3.1) and the expression of the coefficients β's, $\bar{\beta}$'s and the expressions for the T.E.'s have been omitted for brevity; see Usmani [5, Chap. 3, pp. 29–35].

The finite difference equations approximating problem (1.1) can be very conveniently written in matrix form:

$$Ay=\bar{d}, \tag{3.4}$$

where $A=J+h^2DB$ and $\bar{d}=d-h^2Dg$. Here $y=(y_i)$, $d=(d_i)$, $g=(g_i)$ are column vectors, $B=(b_{ij})$, a diagonal matrix such that

$$y_i\simeq y(x_i);\qquad d_1=cy_a-h^2\bar{\beta}_0(f_0y_a+g_0),\qquad d_2=y_a-h^2\beta_0(f_0y_a+g_0),$$

$$d_i = 0 \qquad (i = 3, 4, \ldots, N-2), \qquad d_{N-1} = y_b - h^2\beta_4(f_{N+1}y_b + g_{N+1}),$$

$$d_N = cy_b - h^2\bar{\beta}_0(f_{N+1}y_b + g_{N+1});$$

$$b_{ij} = \begin{cases} f_i, & i = j, \\ 0, & \text{otherwise,} \end{cases}$$

and the matrix D is given by

$$D = \begin{bmatrix} \bar{\beta}_1 & \bar{\beta}_2 & \bar{\beta}_3 & & & & \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 & & & \\ \beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ & & & \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ & & & & \bar{\beta}_3 & \bar{\beta}_2 & \bar{\beta}_1 \end{bmatrix}.$$

If, in addition, in the solution of problem (1.1) $y(x) \in C^8$, then the column vector $Y = (y(x_i))$ (instead of (3.4)) will satisfy the system of equations

$$AY = \bar{d} + T, \tag{3.5}$$

where $T = (t_i)$ is a column vector such that

$$t_i = \begin{cases} \dfrac{c+1}{240} h^6 y^{(6)}(\xi_i), & i = 1, N, \\[2ex] \dfrac{190 - 31c}{60{,}480} h^8 y^{(8)}(\xi_i), & i = 2, 3, \ldots, N-1, \end{cases}$$

provided the parameter c satisfies the inequality (3.3). Further, if we define $E = (e_i)$, the discretization vector, then, on subtracting (3.4) from (3.5), the error equation is obtained in the form

$$AE = T \quad \text{or} \quad E = A^{-1}T. \tag{3.6}$$

Any further analysis now depends on the properties of the matrix $A = J + h^2DB$.

4. ERROR ANALYSIS

As a consequence of Theorem 2.1 it follows that the matrix A is irreducible if

$$\begin{cases} h^2\bar{\beta}_2 f_M < c - 2, \\ h^2\beta_j f_M < c - 2, \qquad j = 1, 3 \quad \text{and} \quad f_M = \max\limits_{a \leqslant x \leqslant b} f(x). \end{cases}$$

Because the $\bar{\beta}_i$'s and β_i's given by (3.2) are all positive for $c \in (2, 14/3]$, the conditions above will automatically be satisfied provided

$$h < [30(c-2)/((3c+26)f_M)]^{1/2}, \qquad c \in (2, 14/3], \tag{4.1}$$

since

$$\beta_1 = \beta_3 > \bar{\beta}_2.$$

Also, by Theorem 2.2, it follows that A is a monotone matrix with $A \geqslant J$; hence $A^{-1} \leqslant J^{-1}$, by (2.5(iv)).

For the derivation of a bound on $|e_m|$ and $||E||$ we will need the following lemmas.

LEMMA 4.1.

$$(c+2)\sum_{n=2}^{m}(-1)^n D_{n-1} = 1 + c + (-1)^m(D_{m-1} + D_m).$$

Proof. The identity above can be directly derived from (2.3). We rewrite Δ.E. (2.3) in the form

$$(-1)^n D_n = (-1)^n c D_{n-1} - (-1)^n D_{n-2}$$

and, on summing up the terms from $n = 2$ to $n = m$, we have

$$\sum_{n=2}^{m}(-1)^n D_n = c\sum_{n=2}^{m}(-1)^n D_{n-1} - \sum_{n=2}^{m}(-1)^n D_{n-2}.$$

On rewriting the preceding equation so that the suffix of D becomes $n-1$ in all terms and noting that $D_1 = c$, we obtain

$$\sum_{n=3}^{m+1}(-1)^{n-1}D_{n-1} = c\sum_{n=2}^{m}(-1)^n D_{n-1} - \sum_{n=1}^{m-1}(-1)^{n-1}D_{n-1},$$

$$-\left[-D_1 + \sum_{n=2}^{m}(-1)^n D_{n-1} + (-1)^{m+1}D_m\right] = c\sum_{n=2}^{m}(-1)^n D_{n-1}$$

$$+\left[-D_0 + \sum_{n=2}^{m}(-1)^n D_{n-1} - (-1)^m D_{m-1}\right],$$

or

$$(c+2)\sum_{n=2}^{m}(-1)^n D_{n-1} = 1 + c + (-1)^m (D_{m-1} + D_m).$$

LEMMA 4.2.

$$(c+2)\sum_{n=m+1}^{N-1}(-1)^n D_{N-n}$$
$$= (-1)^{N+1}(1+c) + (-1)^{m+1}(D_{N-m-1} + D_{N-m}).$$

Lemma 4.2 can easily be proved in the same manner as Lemma 4.1.

LEMMA 4.3.

$$D_N = D_m D_{N-m} - D_{m-1} D_{N-m-1}, \qquad m = 1, 2, \ldots, N.$$

The identity is obviously true for $m = 1$ using the Δ.E. (2.3). Assume it to be true for $m = p - 1$. Then

$$\begin{aligned} D_N &= D_{p-1} D_{N-p+1} - D_{p-2} D_{N-p} \\ &= D_{p-1}(cD_{N-p} - D_{N-p-1}) - (cD_{p-1} - D_p) D_{N-p}, \qquad \text{by (2.3)}, \\ &= D_p D_{N-p} - D_{p-1} D_{N-p-1}. \end{aligned}$$

Thus the identity is also true for $m = p$ and the lemma follows by mathematical induction.

Define

$$\Delta = \frac{(-1)^m D_{N-m} + (-1)^{N-m+1} D_{m-1}}{D_N}. \tag{4.2}$$

LEMMA 4.4.

(i) $\Delta \leqslant \operatorname{sech}[(N+1)\theta/2]$, *equality holds for N odd and $(N+1)/2$ even,*

(ii) $\Delta \geqslant -\operatorname{sech}[(N+1)\theta/2]$, *equality holds for N odd and $(N+1)/2$ odd,*

(iii) $-1 < \Delta < 1$.

Proof. (i) Consider Δ defined by (4.2) as a function of the real variable m with D_m defined by (2.4) for integral values of m. It is easy to verify that Δ is symmetric in the interval $[1, N]$ and attains its maximum for $m = (N+1)/2$. Since m is a positive integer, Δ will attain its maximum for N odd and $(N+1)/2$ even. Thus

$$\Delta \leqslant \frac{2D_{(N-1)/2}}{D_N} = \frac{2\sinh(((N+1)/2)\theta)}{\sinh(N+1)\theta}, \qquad \text{by (2.4),}$$

$$\leqslant \operatorname{sech}\left[\frac{(N+1)\theta}{2}\right].$$

(ii) Similarly, if N is odd and $(N+1)/2$ odd, then

$$\Delta \geqslant -\operatorname{sech}\left[\frac{(N+1)\theta}{2}\right].$$

(iii) $-1 < \Delta < 1$ follows from the fact that $\operatorname{sech}[(N+1)\theta/2] < 1$ for $\theta > 0$, since $2\cosh\theta = c > 2$.

We can now write the error equation (3.6) in the form

$$|E| \leqslant A^{-1}|T| \leqslant J^{-1}|T|$$

so that

$$|e_m| \leqslant \sum a_{mn}|t_n| = |t_1|(a_{m1} + a_{mN}) + |t_2|\left(\sum_{n=2}^{m} a_{mn} + \sum_{n=m+1}^{N-1} a_{mn}\right)$$

$$\leqslant \frac{|t_1|(1-\Delta)}{c+2} + \frac{|t_2|}{c+2}\left[\left(\frac{m(N-m+1)}{2} - 1\right) + (-1)^m \frac{D_{N-m}}{D_N}\right.$$

$$\left. \cdot \sum_{n=2}^{m} (-1)^n D_{n-1} + \frac{(-1)^{-m} D_{m-1}}{D_N} \sum_{n=m+1}^{N-1} (-1)^n D_{N-n}\right],$$

$$\leqslant \frac{|t_1|(1-\Delta)}{c+2} + \frac{|t_2|}{c+2}\left[\left(\frac{m(N-m+1)}{2} - 1\right)\right.$$

$$\left. + \frac{D_m D_{N-m} - D_{m-1} D_{N-m-1}}{(c+2)D_N} + \frac{c+1}{c+2}\Delta\right],$$

by Lemmas 4.1 and 4.2,

$$|e_m| \leqslant \frac{|t_1|(1-\Delta)}{c+2} + \frac{|t_2|}{c+2}\left[\left(\frac{m(N-m+1)}{2} - 1\right)\right.$$

$$+ \frac{1}{c+2} + \frac{c+1}{c+2}\Delta\Bigg], \qquad \text{by Lemma 4.3,} \tag{4.3}$$

and finally

$$\|E\| \leqslant \frac{1}{c+2}\left[\frac{c+1}{240} h^6 \bar{M}_6 \left(1 + \operatorname{sech}\left(\frac{N+1}{2}\theta\right)\right)\right.$$
$$+ \frac{190 - 31c}{60{,}480} h^8 M_8 \left\{\left(\frac{(N+1)^2}{8} - 1\right) + \frac{1}{c+2}\right.$$
$$\left.\left. + \frac{(c+1)\operatorname{sech}(((N+1)/2)\theta)}{c+2}\right\}\right], \qquad \text{by Lemma 4.4,} \tag{4.4}$$

or

$$\|E\| \leqslant \frac{h^6\left[\dfrac{(c+1)\bar{M}_6}{120} + \dfrac{(190-31c)(b-a)^2 M_8}{483{,}840}\right]}{c+2} \tag{4.5}$$

(since $\operatorname{sech}((N+1)\theta/2) < 1$), where

$$\bar{M}_6 = \max|y^{(6)}(x)|, \qquad x \in [(a, x_3) \cup (x_{N-2}, b)]$$

and

$$M_8 = \max|y^{(8)}(x)|, \qquad x \in [a, b].$$

We summarize the preceding results in Theorem 4.4.

Theorem 4.4. *Let $y(x) \in C^8$ be the exact solution of the two-point linear boundary value problem* (1.1) *and let y_n $(n = 1, 2, \ldots, N)$ be the exact solution of the system of linear equations* (3.4) *based on* Δ.E.'s (3.1). *Further, if*

(i) $f(x) \geqslant 0$ *on* $[a, b]$,

(ii) $2 < c \leqslant 14/3$, *inequality* (3.3),

(iii) $h < \left[\dfrac{30(c-2)}{((3c+26)f_M)}\right]^{1/2}$, *see* (4.1),

then

$$\|E\| = 0(h^6),$$

given by (4.5) *or, more precisely, by* (4.4), *neglecting all round-off errors.*

REFERENCES

1 A. R. Aziz and B. E. Hubbard, Bounds for the solution of the Sturm Liouville problem with applications to finite difference methods, *J. SIAM* 12(1964), 163–168.

2 J. H. Bramble and B. E. Hubbard, On a finite difference analogue of an elliptic boundary problem which is neither diagonally dominant nor of nonnegative type, *J. Math. and Phys.* 43(1964), 117–132.

3 C. Froese Fischer and R. A. Usmani, Properties of some tridiagonal matrices and their application to boundary value problems, *SIAM J. Numer. Anal.* 6(1969), 127–142.

4 P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations*, John Wiley, New York, 1961.

5 R. A. Usmani, Numerical solution of boundary value problems in ordinary differential equations, Ph. D. Thesis, University of British Columbia, Vancouver, B.C., March 1967.

6 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

Received May, 1970